Background

The “Victorian Ghost Stories” project arose out of our team leader, Gabi Keane’s, interest in supernatural Victorian literature garnered from previous research performed at the University of Pittsburgh. Our basic goal was to explore short ghost stories written by female English authors during the Victorian Era and uncover some unique insight or claim about how they used language to build suspense. While these stories were consumed mostly by middle class women, they were written by members of the upper class. We became curious as to what “scary words” were used to build suspense and how, if at all, the class or education level of the audience impacted the type of vocabulary these authors used to communicate.

For our readings, we selected five short stories from four different authors: *Walnut-Tree House* and *The Old House in Vauxhall Walk* by Charlotte Riddell; *The Old Nurse’s Story* by Elizabeth Gaskell; *The Shadow in the Corner* by Mary Elizabeth Braddon; and *John Charrington’s Wedding* by Edith Nesbitt. We chose multiple authors to measure for consistency across genre as opposed to one particular artist’s breadth of work. What we’ve come to find is that these stories

Stories by Charlotte Riddell and Elizabeth Gaskell were republished thanks to gothictexts.wordpress.com. *John Charrington’s Wedding* by Edith Nesbitt was available for free via Project Gutenberg. *The Shadow in the Corner* was republished with permission from gothichorrorstories.com. We would like to extend a special thanks to our project overseer, Shannen Davis, as well as Professor David Birnbaum from the University of Pittsburgh for teaching us the tricks of the trade.

Methodology

The “Victorian Ghost Stories” project attempts to answer the following research question: In what ways did female authors use language to build suspense in the Victorian ghost story? Our research topic involved the fields of literature, history, linguistics, and digital research methods. Our approach required textual analysis, XML language tools, and the WordNet lexical database.

For our readings, we selected five short stories from four different authors: *Walnut-Tree House* and *The Old House in Vauxhall Walk* by Charlotte Riddell; *The Old Nurse’s Story* by Elizabeth Gaskell; *The Shadow in the Corner* by Mary Elizabeth Braddon; and *John Charrington’s Wedding* by Edith Nesbitt. This range generated enough data to compare contemporary texts with different writing styles without sacrificing our ability to perform in depth textual analysis on each individual story. Our model is applicable to larger samples of ghost stories, however, should future digital humanists desire to expound upon our research.

After deciding upon our research question and selecting stories, we next performed a close read and analysis of each text to determine our methodology for markup. We decided that the focus of our exploration should be to find “scary words” and rate them according to *scariness*, *part of speech*, and *synset*, so that we could determine not only what scary language was used, but also in what ways its scariness and frequency affected the stories. We accomplished this by compiling a list of 286 unique scary words that we then tagged in each text they occurred. The scariness rating was determined by @scale attributes ranging from 0-3, with the additional rating of “i” for intensifiers like “very” or “quite.” Scary words with a scale of 0 had the potential to ignite fear, but were not so used in context, like the words “black” or “cold.” These words could receive higher ratings when appropriate. Scary words with a scale of 3 were the scariest and therefore always scary, like “horror” or “deadly.” This was the most subjective part of the project, though we gave ourselves parameters in that we drew from an agreed upon list.

Part of speech was measured with attribute @pos, with the available options of verb, noun, adjective, adverb, or modifier. We were curious about the way in which scariness was communicated, and this allowed us to see it at a grammatical, literary level. Is scariness more descriptive, such that the most frequent type of scary words are adjectives? Or is it more action based, such that they are verbs? This data influenced our conclusion, adding another level of specificity to our research question.

Synset is BLANK…

After tagging all of our scary words, we then tracked the frequency and placement of each using XSLT. Frequency was determined by how many times a scary word appeared in a given story. This was done by BLANK. For our visualizations, we selected only words with a scariness rating of 2 or 3. Placement of scary words was used to determine exactly how these authors built suspense. Do more scary words appear in the beginning of the story, or near the end? Are scarier words clustered together? To determine this, we divided the text into five plot points—*introduction*, *protagonist*, *history*, *haunting*, and *resolve*—and represented our data using BLANK to create a circle graph.

These plot points were attributes of the empty element <plot>, which we used as a milestone in our markup. *Introduction* represents the introduction to the story, where the setting and scenario is established. *Protagonist* signifies the introduction of the main character, which has a unique trope in Victorian ghost stories. *History* signifies the moment when a character begins recounting the history of the ghost to the protagonist. *Haunting* signifies the beginning of the ghost’s actual haunting of the protagonist. *Resolve* occurs when the protagonist begins looking for a way to resolve the situation, usually bringing peace to the troubled ghost by righting some wrong that was done to it. Using these elements as milestones, we were able to determine how many scary words each section contained, how scary those words were, and how often they appeared. Visualizations for our findings is shown in detail on our analysis page.